Current Topics in Genome Analysis

Roger H. Reeves, Ph.D.
Department of Physiology
Johns Hopkins University
School of Medicine
rreeves@welch.jhu.edu

COMPARATIVE GENOMICS

Biology may be viewed simplistically as a product of the processes of development, metabolism and aging. These processes are regulated/determined/occur due to an interaction of environmental, stochastic, and genetic factors.
Completion of the initial goals of the genome project will provide maps and sequence of human and model organisms. This information set contains a wealth of information about the genetic contribution to biology. *How can data relevant to the biological goals of the genome project be extracted from this mass of sequence? Perhaps more to the point, what information should be extracted?* Clearly, life scientists want to use this information to understand basic processes of biology in normal and (abnormal) disease states.
A complete, base-perfect human genome sequence provides few *direct* links to human biology. (For example, the "normal" human genome sequence is a tool for disease gene finding, but presumably will contain few or no variants expressed as major Mendelian mutations). Model organisms provide an important part of this critical link in Genomics, as they have throughout the history of Biology.

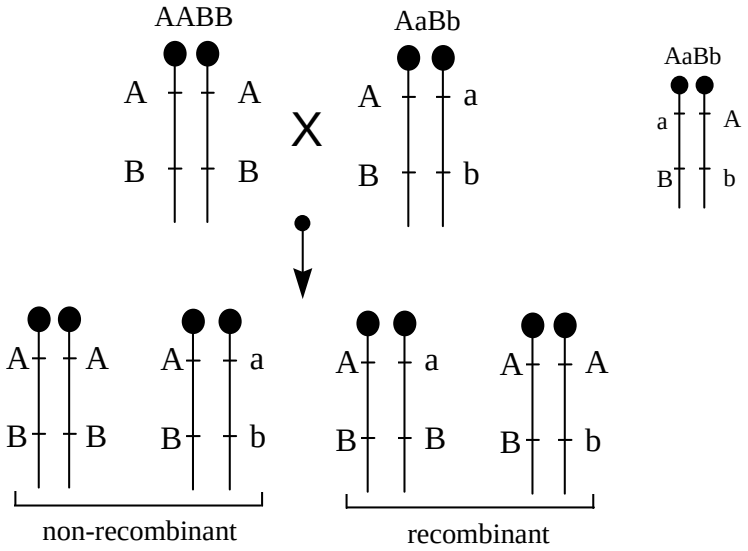
1. Biology (mouse:man)
    - what is phenotype? can mouse and human phenotypes be "the same"? (conserved genetic pathways/ conserved function/ conserved structures)
    - development – metabolism – aging
    - evolution ("development" of species)
2. Learning from evolution
    - comparative sequence
    - comparative structure/ process
3. Why mouse?
    - genetics – *unique genetic structures not found in nature,* nor elsewhere in the laboratory
    - Mouse vs. rat: use the appropriate model system; one is not inherently "better" than the other
    - *ability to create new genotypes and phenotypes*
4. Building custom phenotypes and genotypes
    - mutagenesis
    - transgenesis/ targeted gene modification/ chromosome engineering
5. QTL/ modifiers
    - real genetics
    - special genetic tools
    - examples

In this lecture we will consider:

1. Genetic structures in mice not found in nature.
2. Relating mouse and human: comparative mapping and sequencing
3. Genetics meets genomics: Mapping quantitative trait loci
4. Make your own genotype:
   Transgenesis, gene targeting, chromosome engineering.
5. Make your own phenotype: Mutagenesis.

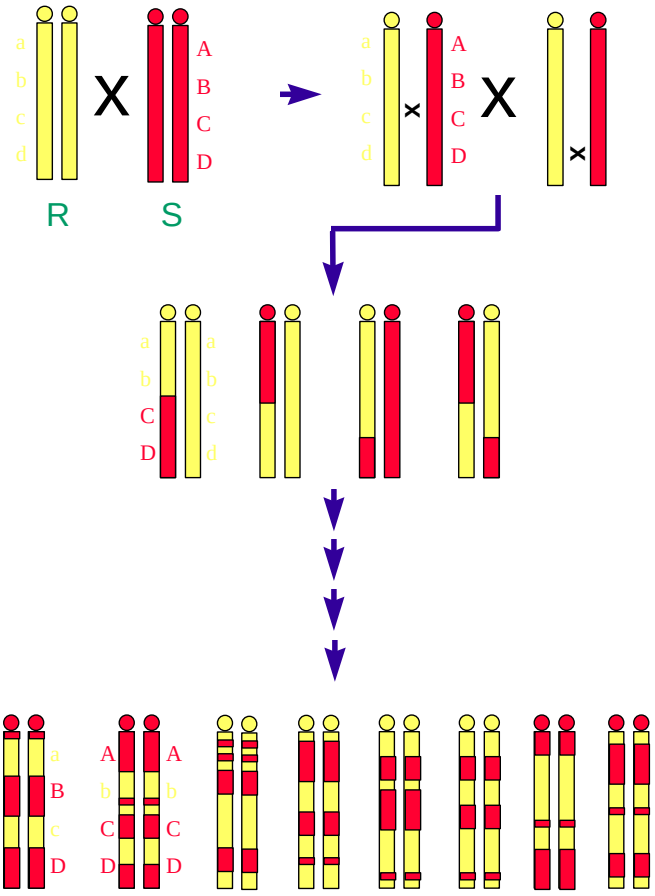
1. Special Genetic Structures in Mice

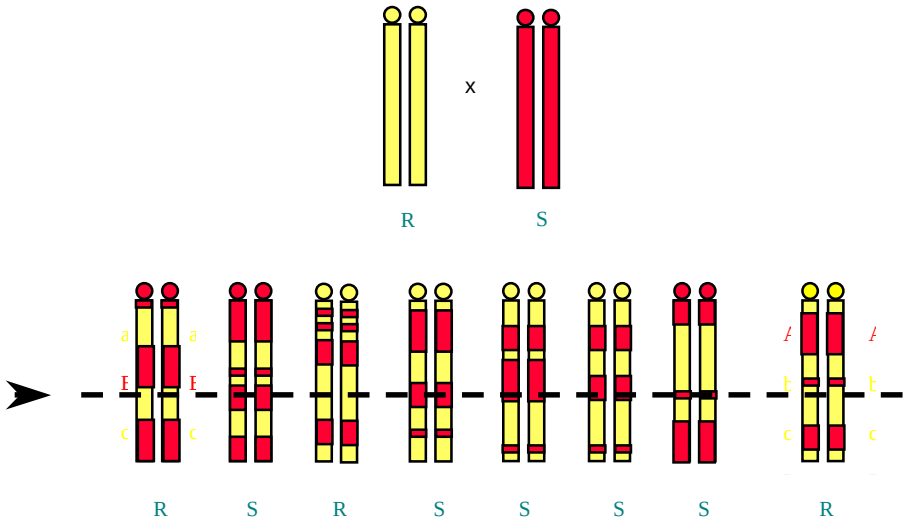
1. Inbred mice – known genetics of individuals and F1's; limited (comprehensively definable) variation in crosses in further generations. Practical benefits of short generation time, compact housing, macro-environment highly controlled in SPF colonies.



Because phase is known, recombinants are unequivocally distinguished from non-recombinants. Individuals in subsequent generations vary genetically, but a maximum of two alleles can be present in at any given locus.

2. Recombinant inbred mouse strains – resource for genotype and phenotype.





Benefits of RI strains:

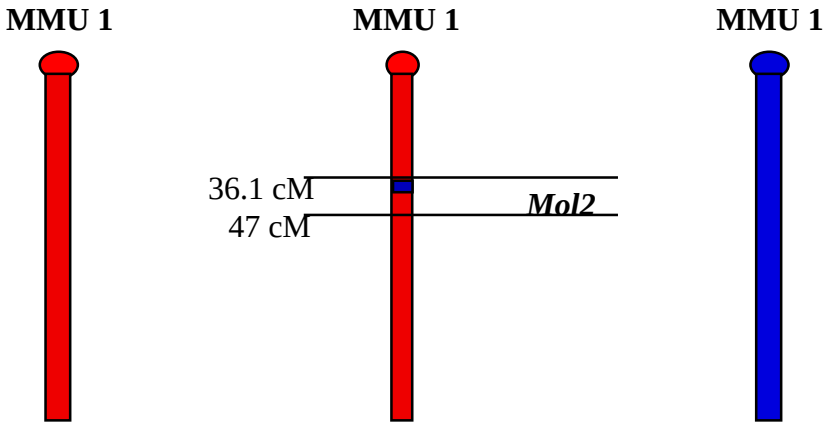
- "Pre-genome scan"
- Reassay the "same" individual many times (find
- true mean and deviation for variable traits)
- Highly beneficial for quantitative traits
- Stock of genetic variation (recombinant congenic
- mice)

Limitations of RIs:

- Small strain sets have limited statistical power.
- Mapping is relatively low resolution on the first pass.

3. Congenic mice can be used to confirm and refine localizations made with low confidence initially.

Independent confirmation of *Mol2* on Chr1



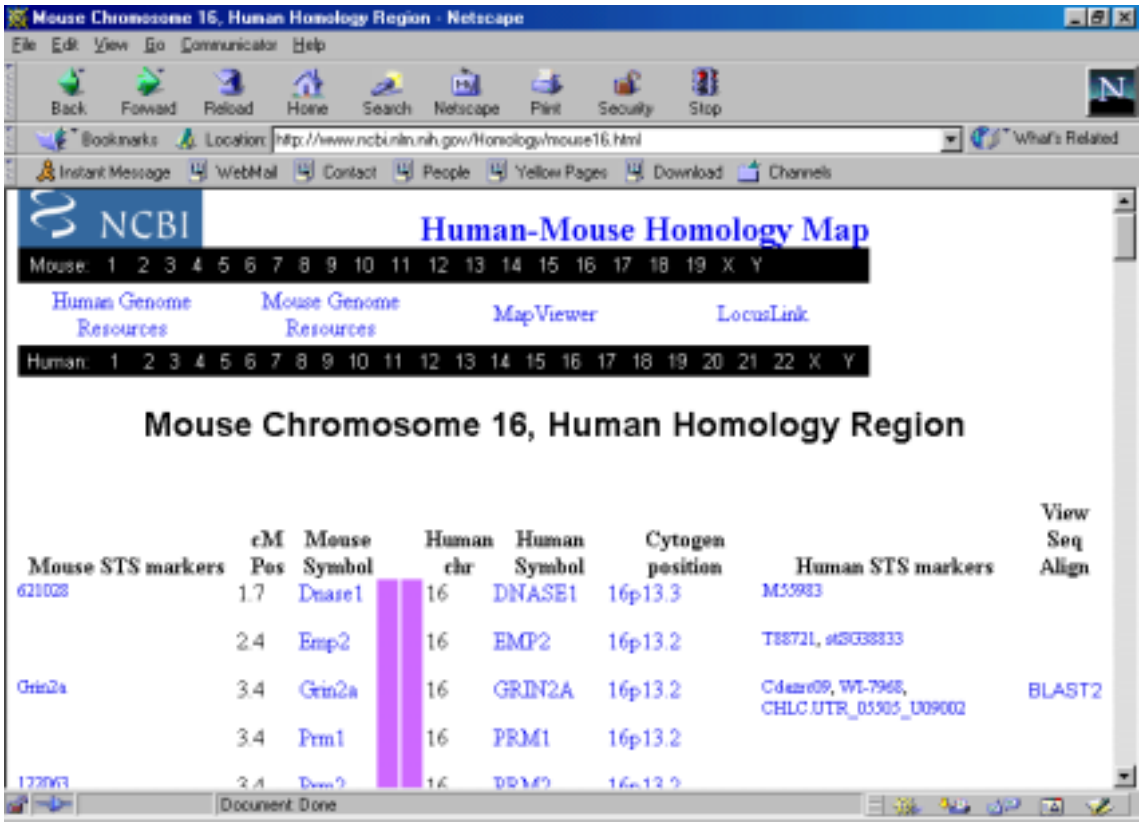
STRAIN	BALB/cAnFCR	C.D2-VIL6	DBA/2N
RESPONSE	0%	79% + 33% (n=6)	100%
	Low Responder	High Responder	High Responder

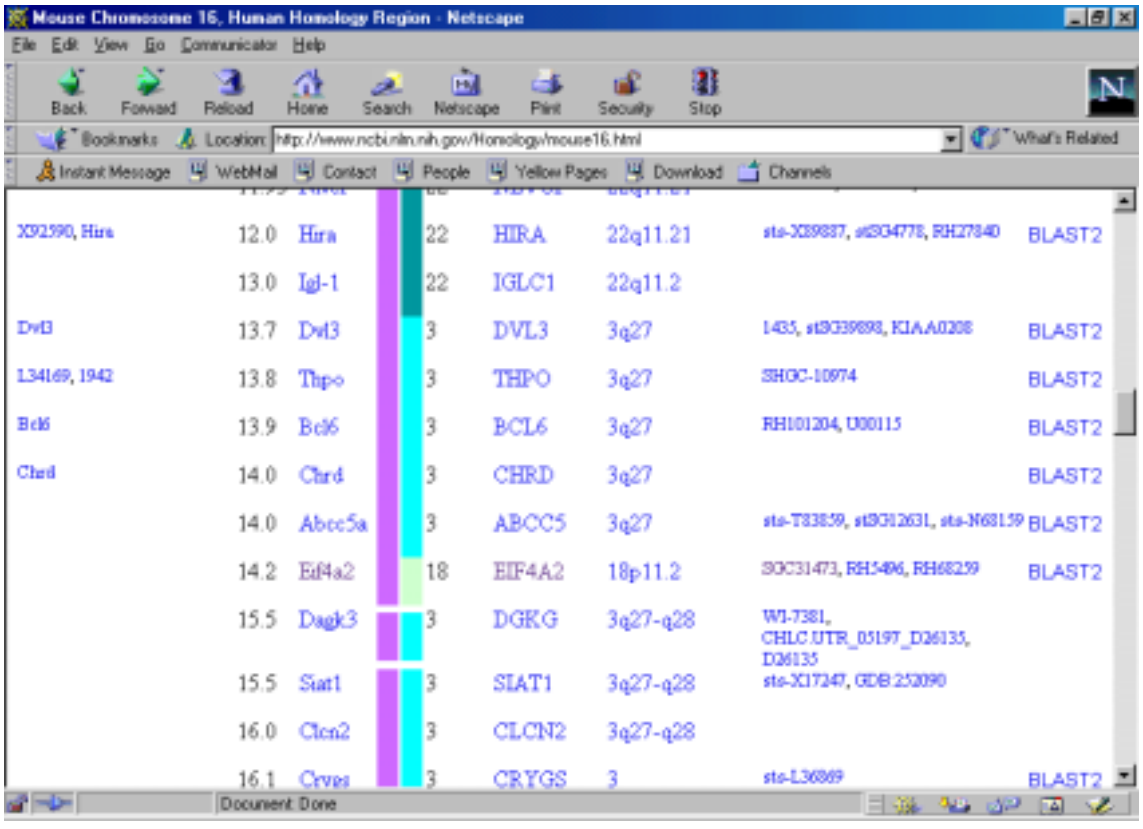
Matesic, LE, EL Niemitz, A De Maio, and RH Reeves. 2000. Quantitative trait loci modulate neutrophil infiltration in the liver during LPS-induced inflammation. FASEB Journal (November, 2000)3. QTL vs. Mutagenesis

Comparative Genetics, Mouse vs. Human

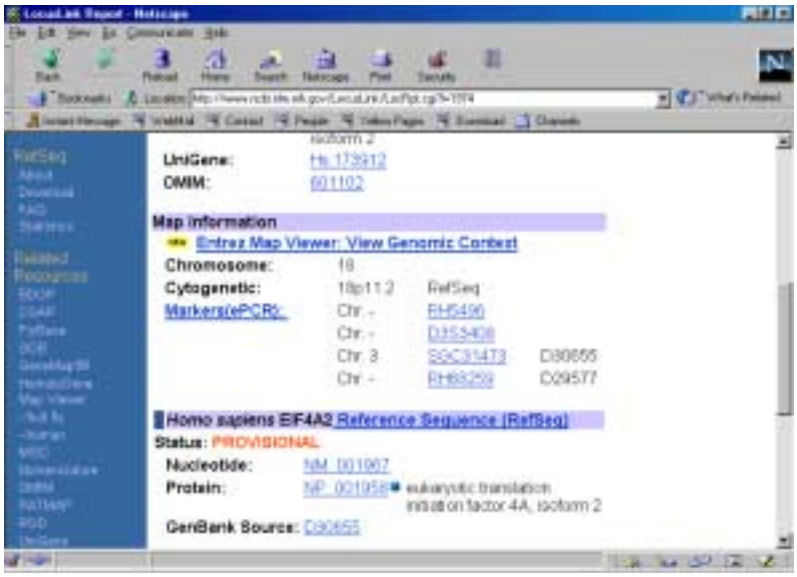
Comparative mapping

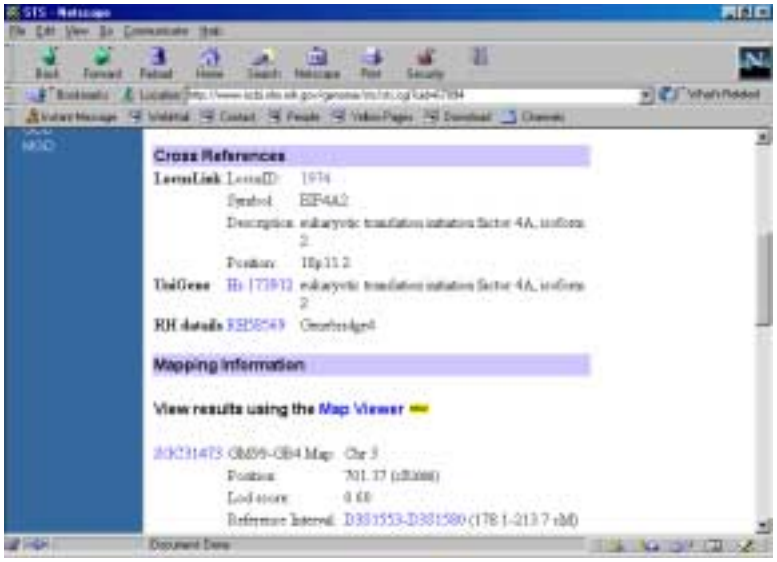
http://www.ncbi.nlm.nih.gov/Homology/
http://www.informatics.jax.org/



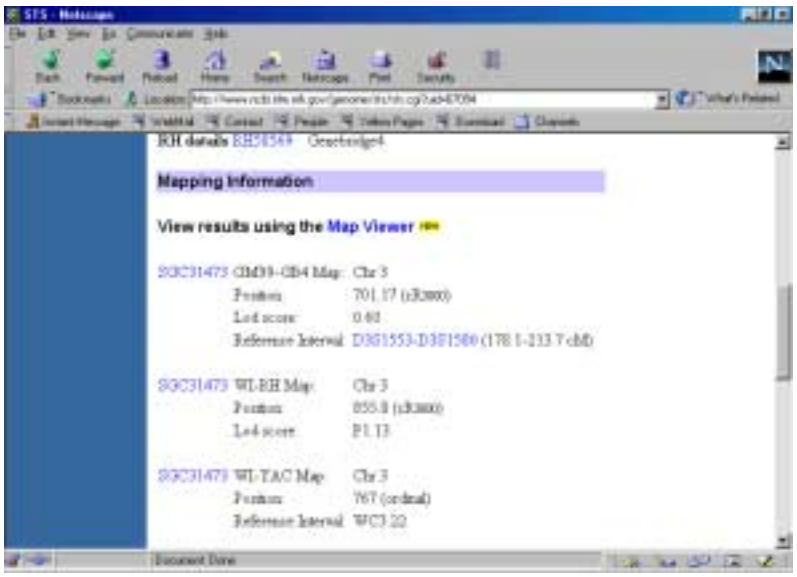


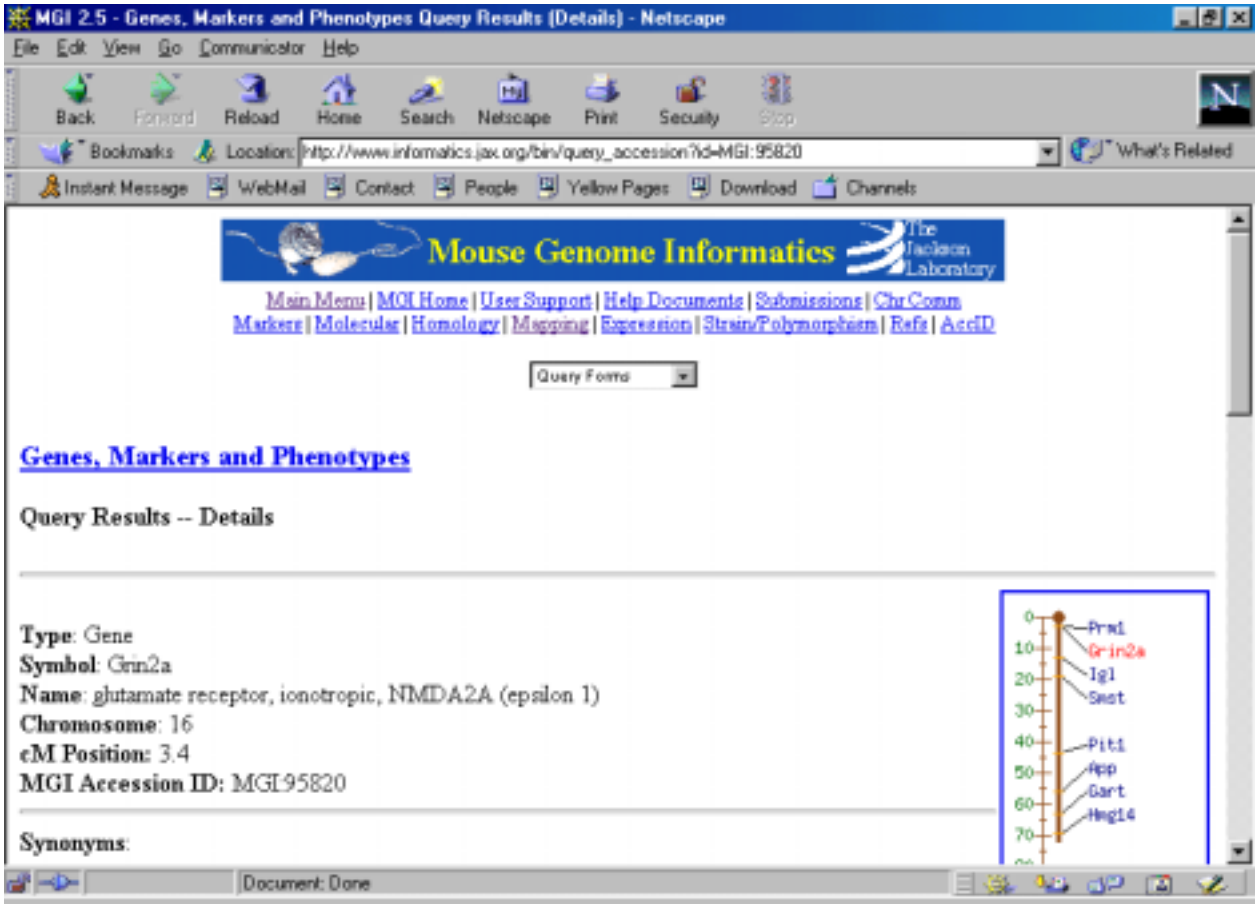
LOCUS LINK





UNI-STS





Comparative sequence analysis

Comparative Genome Analysis Tools (CGAT) can be downloaded from http://inertia.bs.jhmi.edu/roger/CGAT/CGAT.html. See also Rosetta and Glass at http://www.theory.lcs.mit.edu/crossspecies.

CGAT: Lund J, Chen F, Hua A, Roe B, Budarf M, Emanuel BS, Reeves RH. Comparative sequence analysis of 634 kb of the mouse chromosome 16 region of conserved synteny with the human velocardiofacial syndrome region on chromosome 22q11.2. Genomics. 2000 Feb 1;63(3):374-83.

Also:

Bouck JB, Metzker ML, Gibbs RA. Shotgun sample sequence comparisons between mouse and human genomes. Nat Genet. 2000 May;25(1):31-3.

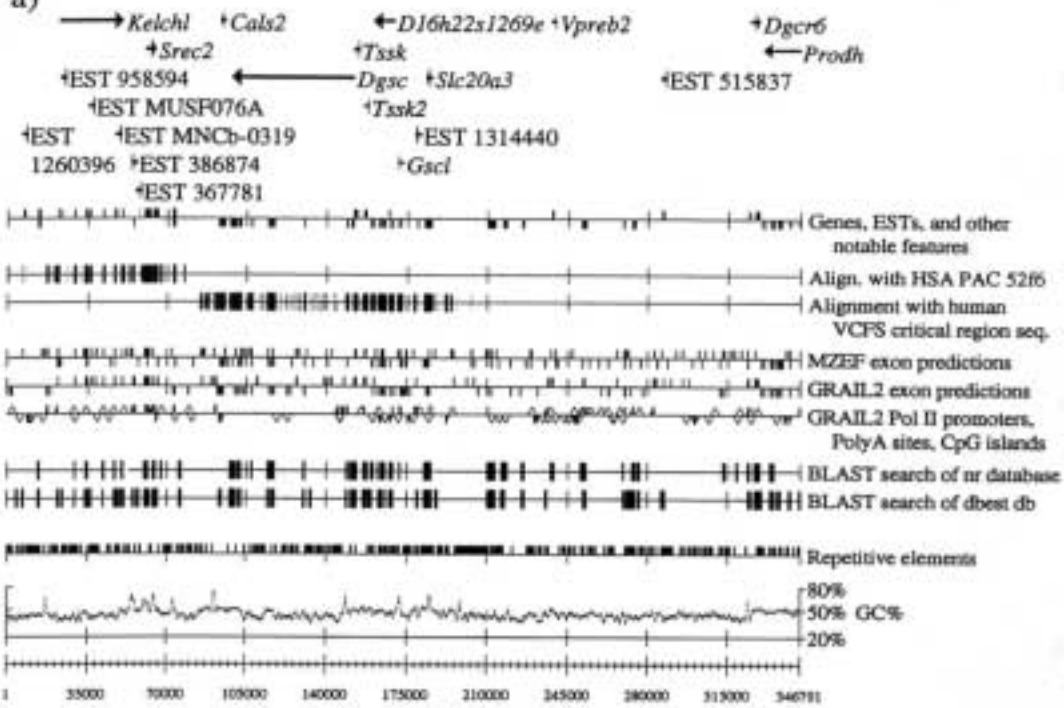
Rosetta: Batzoglou S, Pachter L, Mesirov JP, Berger B, Lander ES. Human and mouse gene structure: comparative analysis and application to exon prediction. Genome Res. 2000 Jul;10(7):950-8.

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W. PipMaker--a web server for aligning two genomic DNA sequences. Genome Res. 2000 Apr;10(4):577-86.

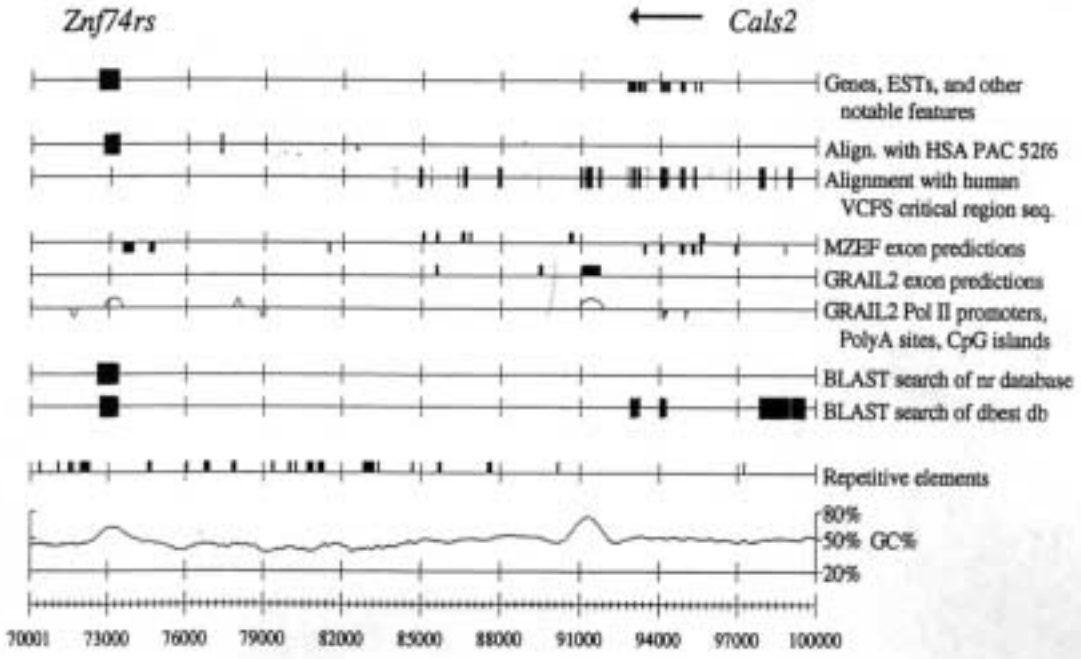
Reeves, Comparative Genomics

Comparative sequence analysis is invaluable for finding and defining the boundaries of genes and other conserved features. Let evolution work for you! 80 million years of conservation can't be wrong!

a)



b)



Comparative sequence +/- algorithmic predictions

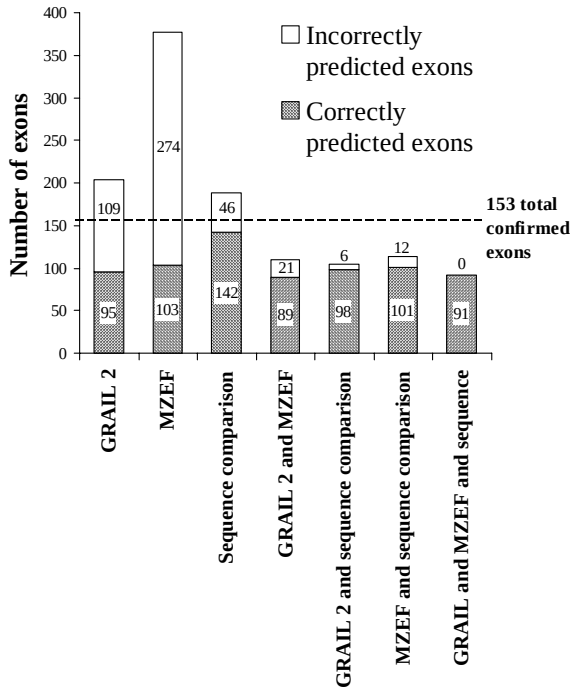


Fig. 5. Comparison of exon prediction approaches from a 500kb segment of conserved synteny between human Chr 22 and mouse Chr 16 containing 153 exons from 18 genes (Lund et al.,). Sequence comparison identifies only conserved segments, eliminating the false positive predictions in algorithmic approaches. In conjunction, these approaches identify at least a portion of every gene with no false positive predictions.

Mapping Quantitative Trait Loci (QTL)

1. Genomic genetics- why is the genome complex?

How does it work (how do genes in many variant combinations in a population act together) to maintain homeostasis in the face of infinitely complex and dynamic challenges from the environment?

2. Plants have the lead. The tomato as a model for mammalian genetics.

Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science. 2000. 289(5476):85-8.

3. Gene interactions

		Genotype at *Hpi2*, Chromosome 5							
		A/A		A/B		B/B		Totals	
Genotype at *Hpi1*, Chromosome 13	A/A	33.5 (9)	± 4.6	35.6 (12)	± 4.8	35.6 (8)	± 6.9	35.0	± 3.0
	A/B	28.9 (11)	± 5.0	35.7 (40)	± 3.0	37.8 (11)	± 4.8	34.9	± 2.3
	B/B[b]	42.5 (2)	± 4.1	44.7 (14)	± 5.3	69.9[c] (11)	± 5.5	54.8	± 4.3
	Totals	32.0	± 3.2	37.6	± 2.3	49.0	± 4.3	39.5	± 1.9

[a] Avg. number of PMN per h.p.f. ± s.e. are given for (n) animals of each genotype class.[b] Mice with a B/B genotype at *Hpi1* showed significantly higher PMN infiltration values than other *Hpi1* genotypes (p=1.22 X 10-4, t-test assuming unequal variance)[c] Mice with a B/B genotype at both Hpi1 and Hpi2 showed significantly higher PMN infiltration than other genotype classes (p=7.83X10-5, t-test assuming unequal variance)

Matesic, LE, EL Niemitz, A De Maio, and RH Reeves. 2000. Quantitative trait loci modulate neutrophil infiltration in the liver during LPS-induced inflammation. FASEB Journal (November, 2000)3. QTL vs. Mutagenesis

4. QTL mapping in human beings (behavioral disorders, neuropsychiatric disease, diabetes, metabolic regulation).

- Very high marker density
- Very large numbers of patients
- Precise diagnosis

4. Transgenesis/ gene-targeting/ chromosome engineering

1. "Knockouts" (null alleles)

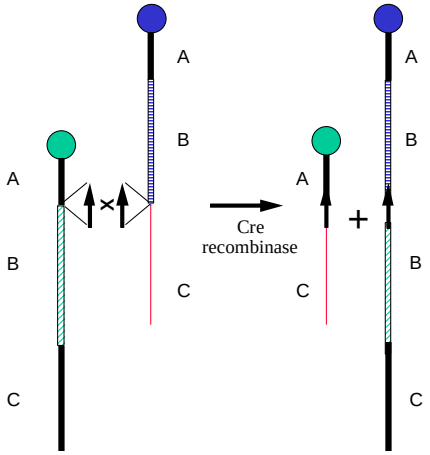
2. "Knock-ins" (mutations, reporters), tissue-targeted and conditional mutations

Shin MK, Levorse JM, Ingram RS, Tilghman SM. The temporal requirement for endothelin receptor-B signalling during neural crest development. Nature. 1999 Dec 2;402(6761):496-501.

3. Chromosome engineering

Ramirez-Solis R, Liu P, Bradley A. Chromosome engineering in mice. Nature. 1995 378(6558):720-4.

Cre/lox-mediated chromosomal translocation



4. Whole genome gene deletion strategies

Zheng B, Mills AA, Bradley A. A system for rapid generation of coat color-tagged knockouts and defined chromosomal rearrangements in mice. Nucleic Acids Res. 1999 27(11):2354-60.

Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, Sands AT. Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature. 1998 392(6676):608-11.

xerox Sands diagram

5. Mapping Genome Function: Creating Phenotypes using Mutagenesis

Mutagenesis provides a means of generating new phenotypes in mouse.

Justice MJ, Zheng B, Woychik RP, Bradley A. Using targeted large deletions and high-efficiency N-ethyl-N-nitrosourea mutagenesis for functional analyses of the mammalian genome. Methods. 1997 Dec;13(4):423-36. Review.
MJ Justice in IJ Jackson and CM Abbott, Mouse Genetics and Transgencis: A Practical Approach. 2000. Oxford University Press, 299 pp.

1. Sources of mutations
    spontaneous, E-5, all types of mutations;
    radiation, frequency is dose dependent, primarily chromosomal rearrangement;
    chemical, ENU is highest giving point mutations at a frequency of 1/600 gametes per locus at some loci
2. Screens
    specific locus test
    MutaMouse/ Big Blue
    SHIRPA
    special targeted screens
    dominant vs. recessive (1^{st} vs. 3^{rd} generation)
    in combination with deletion (recessives in first generation)

3. targets/ mutation types
    visible single gene dom. or recessive
    allelic series
    biochemical pathway
    sensitization (Shedlovsky A, McDonald JD, Symula D, Dove WF. Mouse models of human phenylketonuria. Genetics. 1993 Aug;134(4):1205-10.)

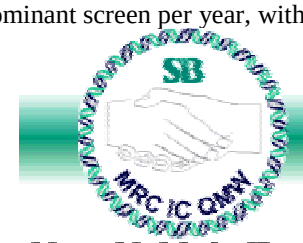
4. mode of action, transfer ethyl group to a number of residues on all four nucleotides, including ethylation of phosphate groups that leads to mispairing;
    most frequent in mouse are AT -> TA (transversion) and AT -> GC (transition), but specific frequencies are locus specific
    ENU affects primarily spermatogonia (stem cells) – freq. in sperm and in females are low

5. Breeding schemes:
a. balancer;
b. recessive over deletion;
c. modifier (dominant mutation modifies another mutation)
d. sensitization (recessive mutations in genes that interact in a pathway/ allelic non-complementation)

6. Large centers, see Trans-NIH Mouse initiative http://www.nih.gov/science/models/mouse/index.html

Mouse Genome Center, Harwell http://www.mgu.har.mrc.ac.uk/mutabase/
ENU Mutagenesis Programme

We are engaged in a major ENU mutagenesis programme that incorporates the systematic and semi-quantitative screening protocol - SHIRPA. Spanning the next three years, a genome-wide screen for dominant mutations will be carried out. BALB/c males mutagenised with 160-200 mg/kg ENU are being mated to C3H females and 40,000 of their F1 progeny will be characterised using the SHIRPA protocol. From screening results to date, it appears that approximately 1% of the F1 population represent inherited mutations. This number should increase as additional screens are added to the programme. It is planned to map around 50 mutations from this dominant screen per year, with a proportion undergoing more in-depth characterisation.



ENU Mutagenesis Programme

Mouse Models for Human Disease

Inherited Mutations Identified So Far

Bottom Home Email the Mutabase team

SHIRPADATA	MUTANTLINE#	MOUSE #	SEX		ANOMALY
SHIRPADATA	6	4.2a	f		Bpa like
SHIRPADATA	15	30.2 i	m		Small (75% at birth and weaning, smaller in progeny), left ear lower, coat slightly darker, (one or more of these anomalies inherited). SHIRPA . Cross to BALBc for mapping
SHIRPADATA	22	3.2a	f		Stripes (anaemia, cataract, short head, males don't survive past weaning ?). Sex linked anaemia.
SHIRPADATA	25	61.7 e	f		Circling, head bobbing, cataract in rt eye, some progeny with small head/snout.
SHIRPADATA	26	61.6 a	f		White feet, head spot and belly spot (some progeny with short tails, smalls @ b and w). Limb tone in shirpa (operator sensitive?)

Top Home Email the Mutabase team

German ENU Mutagenesis Center http://www.gsf.de/isg/groups/enu-mouse.html

We plan to use two different strategies in order to screen for new mouse mutants: A dominant and a recessive screen. In both cases, male mice are injected with ENU and then mated to females in order to produce F1 founders. These F1 mice can either be analyzed directly for dominant mutations or bred further to subsequently study recessive phenotypes. Very large numbers of mice can be analyzed in a dominant F1 screen. In this case, all F1 mice are screened for phenotypic abnormalities. If the animals might die during the screening procedure, F2 mice are produced and analyzed. F1 mice are preserved for breeding the potential mutants. The screen for recessive mutations will involve two generations of breeding. From F1 founder males, F2 female offspring are raised, half of which are heterozygous for the newly induced mutations. Backcrossing F2 to the F1-founder male or intercrossing the F2 is then carried out to identify recessive mutant phenotypes among the F3 offspring.